Eighth International Conference on Construction in the 21st Century (CITC-VIII) "Changing the Field: Recent Developments for Future Engineering and Construction" May 27-30, 2015, Thessaloniki, Greece

The Validation of a Sexual Discrimination Measure for Construction

Valerie Francis
Faculty of Architecture, Building and Planning
vfrancis@unimelb.edu.au

Abstract

Research in construction field indicates women's lack of career achievement is related to exclusionary and discriminatory practices within their organisations. However, few measures of this gender equity have been developed, with most typically assuming a strong gender bias. This study developed and validated a new measure for professional men and women in construction which has been described within two separate papers. The first paper (Francis, 2015) presented the development of a new 20-item trial measure using an iterative process and the assistance of a ten person expert panel. This paper is the second and describes the final part of the study presenting the validation of the new measure. Nine items, including two relating to sexual harassment, had to be removed as they could not be considered as continuous or were non parametric. Because of the exclusion of the sexual harassment items the final scale was a measure of sexual discrimination. To establish which of the 11-items should be retained, an Exploratory Factor Analysis (EFA) was conducted using a sample of 217 professional women in construction. The final scale included 5-items and was verified using a Confirmatory Factor Analysis (CFA) using a sample of 239 construction women. The Cronbach Alpha for the scale was found to be .81, indicating a relatively high level of reliability. The research involved important research lessons relating to issues such as the response format, impact of increasing gender neutrality and the reverse scoring of items. It is hoped this study will be useful to other construction researchers contemplating scale development.

Keywords

Construction, scale development, gender inequity, sexual discrimination

1. Introduction

Gender inequity can be considered to involve two subconstructs namely; sexual discrimination, and sexual harassment. It has long been identified as a barrier to women's career progression in construction and known to diminish their workplace experiences (e.g. Gale, 1994; Dainty, et al., 2001; Watts, 2007). This paper is the second of two describing a study involving the development and validation of a new measure of gender equity. This scale is required in order that further research can be conducted to investigate both the extent and effect of inequity on women's careers in construction. This paper outlines the process used to validate a new measure. The final scale contained one factor, sexual discrimination. The reasons for this are explained and some important lessons learnt are summarised.

2. Development of the trail scale

As outlined in Francis (2015) the 30-item Women Workplace Culture Questionnaire (Bergman & Hallberg, 2002) was used as a base from which to develop the new trial measure. This scale had both validity and reliability issues due to the wording of the items, the focus on issues outside the organisation, the lack of consistency within the response format and the gender biased nature of the item wording. A copy of the original scale can be found in Table 1 of Francis (2015).

3. The trial measure

The trial measure had 20-items (refer Table 1), a unique 5-point response format with two subconstructs expected relating to sexual discrimination and sexual harassment. Items e and k of this trial scale were related to harassment and the remaining items related to discrimination. The item stem used was "*When you think about the experiences of professional and managerial staff in your organisation please select the number you consider best reflects what occurs in your workplace.*"

Table 1: Trial scale (Francis, 2015)

Anchor at 1 (on 5-point scale)
a. Men receive high levels of organisational support.
b. Women's views, comments and suggestions are often ignored in meetings.
c. Men have to be highly accomplished in their work in order to be promoted.
d. Women have opportunities to gain professional development.
e. Men are subjected to unwelcome conscious body contact
f. Women are excluded from important work-related conversations.
g. Men seem to "fit in" well.
h. Women work on high profile or challenging projects.
i. Men are successful in obtaining fair remuneration.
j. Women receive unfair judgements of their work performance.
k. Men receive unwelcome sexual comments at work
l. Women enjoy good employment security.
m. Men have a mentor who has a senior role in the organisation
n. Women don't receive enough support to manage their work and domestic responsibilities
o. Men are allocated roles that allow for little career advancement
p. Women find their contribution is highly regarded.
q. Men are allocated roles with lower levels of responsibility
r. Women are included in informal networks
s. Men are considered for promotional opportunities.
t. Women are in mainstream roles

Note: Anchor at 3 was "All professional and managerial staff are treated equally in this matter". Anchor at 5 was reverse of anchor at 1. (e.g. anchor at 5 for a. was Women receive high levels of organisational support).

4. Validation of the measure

To validate a scale two separate samples are generally required, one in which to conduct an Exploratory Factor Analysis (EFA) and one a Confirmatory Factor Analysis (CFA). This two-step procedure is recommended within scale development research (De Vellis, 2003; Worthington and Whittaker, 2006). In this study a sample of 456 professional and managerial women in the Australian construction industry were used to validate the measure. As time did not permit two separate data collections, SPSS ver 20 was used to randomly select 50% of the cases from the main data file. This resulted in a data file with 217 participants, and their responses to the trial scale and used for the EFA. Another file was generated with the remaining 239 participant details and their responses used in the CFA.

Therefore, scale validation had several steps. Initially an assessment of the data was made and then the sample split. The EFA was conducted using the dimension reduction feature in SPSS. The CFA was undertaken by testing the model using structural equation modelling

(SEM) within AMOS. The EFA was conducted to ascertain if the two factor structure expected from the Bergman and Hallberg (2002) scale had been retained and establish which of the items should be retained. As noted, two factors were expected but unfortunately a third item for second factor of sexual harassment had not been included. A minimum of three items are preferred for each factor (Field, 2009). This model was then verified using a CFA.

4.1 Normality tests

Before the EFA and CFA could be conducted normality of the data within the whole sample was assessed. Due to the nature of the sample (all women) it was expected that items could be skewed and this was found to be the case. Unfortunately this also resulted in items b, e, f, j, k and n only being answered on 4 of the 5 possible responses. As at least five ordered categories are required in order for data to be treated as continuous (e.g. Bollen, 1989; Babakus, Ferguson & Jöreskog, 1987) so these items had to be removed from further analyses. This unfortunately also resulted in the elimination of second factor namely sexual harassment. From this point it was apparent that a measure of sexual discrimination, rather than gender equity, was being developed.

The 14 remaining items (a, c, d, g, h, i, l, m, o, p, q, r, s, t) were then assessed for normality. Muthén and Kaplan (1985) consider variables with univariate skewnesss and kurtosis in the range of -1 to +1 as adequate. Finney and DiStefano (2006) in a review of previous studies considering the impact of normality on ML based results found that issues occur when univariate skewness approaches 2 and univariate kurtosis approaches 7. For this reason the values were reviewed and it was considered appropriate to remove items with skewness of greater than 1 (m) and kurtosis of greater than 2 (d and l) from further analysis. (Refer Table 2).

Table 2: Descriptive and normality statistics.

	Mean	Std. Dev	Skew -ness	Std. Error	Kurtosis	Std. Error
a. high organisational support	2.39	.872	-.063	.114	-.110	.228
c. accomplished in work in order to be promoted (rev)	2.20	.947	.164	.114	-.639	.228
d. opportunities to gain prof developm't (rev) *	2.88	.739	-.297	.114	**2.205**	.228
g. "fit in" well	2.39	.915	.009	.114	-.159	.228
h. Work on high profile or challenging projects (rev)	2.64	.866	-.193	.114	.540	.228
i. successful in obtaining fair remuneration	2.25	.837	-.109	.114	-.607	.228
l. enjoy good employment security (rev) *	2.90	.617	-.892	.114	**5.034**	.228
m. have a mentor in senior role in the organisation *	2.74	.711	**-1.013**	.114	2.045	.228
o. allocated roles with little career advancement (rev)	2.71	.699	-.762	.114	1.486	.228
p. contribution is highly regarded (rev)	2.76	.712	-.569	.114	1.884	.228
q. allocated roles lower levels of responsibility (rev)	2.51	.726	-.870	.114	-.047	.228
r. included in informal networks (rev)	2.53	.866	-.244	.114	-.056	.228
s. considered for promotional opportunities.	2.50	.750	-.549	.114	.353	.228
t. in mainstream roles (rev)	3.21	.925	.047	.114	.451	.228

N= 456, Min = 1.0 and Maximum = 5.0. * item has normality issues

In addition to skewness and kurtosis, the Kolmogorov-Smirnov test was also conducted. The results of the Kolmogorov-Smirnov were significantly non normal for each item, D (456) = 0.248 to 0.446, $p < .05$. However, as Field (2009; 144) notes a limitation of this test is that in large sample sizes it is "very easy to get a significant results from small deviations from normality". Field suggests plotting data in order to make an informed decision about the extent of non-normality. This was undertaken and the results indicated the remaining 11 items (a, c, g, h, i, o, p, q, r, s, and t) were reasonably normally distributed which was supported by the normal probably plots (Normal Q-Q Plot – not included due to page restrictions). The EFA was then conducted using the 11 items within the sample of 217 participants.

4.2. Exploratory Factor Analysis (EFA)

4.2.1 Sample size, factor loadings, communalities and KMO

It is important to check the sample is adequate and establish some statistical criteria. While Comrey (1973) indicated that factor analyses require a sample of at least 300, Gorsuch (1983) recommends 5 to 10 participants per item. Velicer and Fava (1998) however consider less than 3 participants per item to be inadequate. Tabachnick and Fidell (2007) indicate that these general rules can be misleading. Guadagnoli and Velicer (1988) argue that the most reliable factor solution depends not only on the sample size but also the factor loadings. They conclude that when there are at least 4 items per factor and loading of greater |.6| then it is reliable regardless of sample size, however for factors with ten or more loadings greater than |.4| are reliable provided the sample is greater than 150. Worthington and Whittaker (2006) indicate that samples of 150 to 200 are adequate providing communalities of higher than 0.5 and (with 10 participants per item) factor loadings at approximately |.4|. Therefore, the sample of 217 was considered to be adequate for the 11 items.

Worthington and Whittaker (2006) also suggest deleting items with factor loading of less than |.32| or cross loading less than .15 difference from the item's highest factor loading. Communalities as suggested by Worthington and Whittaker (2006) need to be greater than 0.5 for an item to be retained if the sample size is less than 300. Field (2009) notes that with a sample of less than 100 communalities of 0.6 and above are required, however for samples of 100 to 200 communalities in the 0.5 range are adequate provided there are relatively few factors. Furthermore the Kaiser-Meyer-Olkin (KMO) measure will be calculated. A KMO close to 1 indicates distinct and reliable factors (Field, 2009). Hutcheson and Sofroniou (1999) indicate values between 0.7 and 0.8 as good and 0.8 and 0.9 as very good. A KMO over 0.8 will be considered adequate. Therefore, to ensure the scale items selected were robust only items with factor loading of |.5| and more will be retained and any secondary loading (if applicable) would be be less than |.3|. Communalities of less than 0.5 with the current sample would be considered problematic. A KMO of greater than 0.8 would also be required.

4.2.2 Extraction and Rotation Methods

An exploratory factor analysis using common-factors analysis (principal-axis factoring) with oblique rotation (Direct Oblimin) was conducted on the 11 remaining items. Common-factors analysis (FA) was selected over principal-components analysis (PCA) as its purpose is more closely aligned with the development of new scales (Worthington and Whittaker, 2009). Gorsuch (1997) indicates that of the two best known techniques that use FA principal-axis factoring have slightly less problems than maximum-likelihood extractions, however maximum-likelihood has been shown to be superior as a CFA technique prior to the current use of SEM (Tabachnick & Fidell, 2007). For this reason FA principal-axis factoring was adopted in the EFA and ML in the CFA. With regard to the choice of rotation method, an oblique rather than orthogonal rotation, was selected. This is due to the fact that an orthogonal rotation is typically chosen, when it is known or assumed that the underlying factors are correlated (Thompson, 2004). In addition, orthogonal rotation of correlated factors tends to over estimate loadings thus leading to the retention or deletion of items and a factor structure which may not be replicated within the CFA (Loehlin, 1998). It was expected that if more than one factor was found that they would be correlated. Direct Oblimin is a method for oblique (non orthogonal) rotation within SPSS.

4.2.3 EFA Analyses

The initial EFA used FA principal-axis factoring, direct oblimin rotation and included the 11 items, seven of which were reversed coded. The Kaiser-Meyer-Olkin measure verified the sampling adequacy for this analysis and was .883, which is well within the limits as indicated earlier. However, the communalities for item t were very low at 0.108 so it was decided to remove this item and rerun the analysis with the 10 remaining items. The initial EFA was then rerun using FA principal-axis factoring, direct oblimin rotation and included the 10 items, six of which were reversed coded. The Kaiser-Meyer-Olkin measure of this analyses was .887. The correlation matrix (refer Table 3) was examined and none were greater than .9 indicating multicollinearity of the data was not an issue. The correlations between the items ranged from .203 to .546 and low correlations related to item h. An inspection of the diagonal values on the anti-image correlation matrix showed values greater than .821 which exceeds the minimum of .5 (Field, 2009).

Table 3: Pearson Correlations

	a.	c.	g.	h.	i.	o.	p.	q.	r.	s.
a. high organisational support	1.000	.382	.510	.240	.455	.313	.325	.456	.421	.546
c. accomplished in work in order to be promoted (rev)	.382	1.000	.357	.300	.347	.459	.346	.392	.308	.492
g. "fit in" well	.510	.357	1.000	.319	.355	.411	.362	.410	.569	.518
h. Work on high profile or challenging projects.(rev)	.240	.300	.319	1.000	.203	.404	.244	.313	.367	.419
i. successful in obtaining fair remuneration	.455	.347	.355	.203	1.000	.329	.283	.440	.259	.571
o. allocated roles with little career advancement (rev)	.313	.459	.411	.404	.329	1.000	.431	.486	.321	.533
p. contribution is highly regarded.(rev)	.325	.346	.362	.244	.283	.431	1.000	.417	.442	.414
q. allocated roles with lower levels of resp (rev)	.456	.392	.410	.313	.440	.486	.417	1.000	.383	.526
r. included in informal networks (rev)	.421	.308	.569	.367	.259	.321	.442	.383	1.000	.361
s. considered for promotional opportunities.	.546	.492	.518	.419	.571	.533	.414	.526	.361	1.000

As recommended the scree plot was examined to determine the number of factors to be retained (Tabachnick & Fidell, 2007; Thompson, 2004). This is shown in Figure 1. The solution (based on Kaiser's criterion of eigenvalues of more than 1 (Kaiser, 1958)) suggests one factor that accounted in total for 40.12% of the variance.

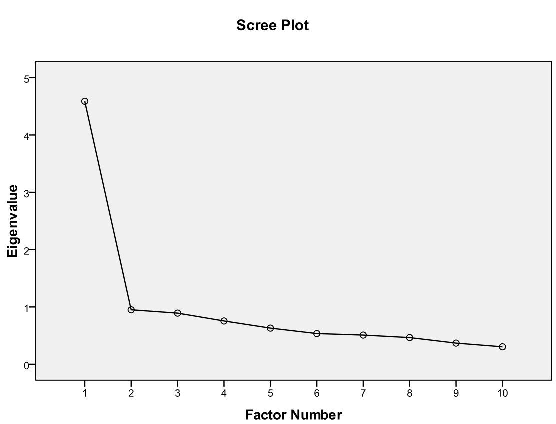

Figure 1: Scree plot of initial Exploratory Factor Analysis (EFA) using 11 items

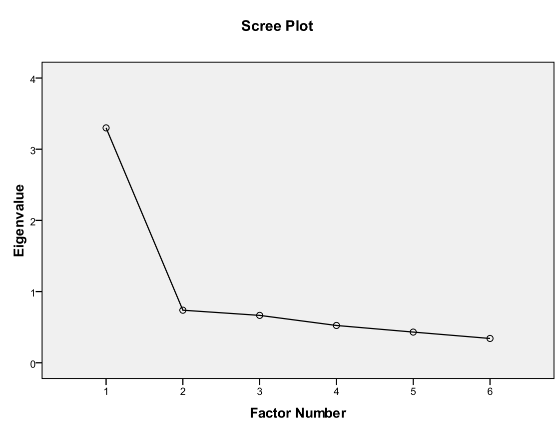

Figure 2: Scree plot of final Exploratory Factor Analysis (EFA) using 5 items

The communalities of the items however, were examined and nine out of the ten items were found to be less than .5. The lowest at .237 was for item h so this was removed and the EFA conducted again. This process was repeated, removing the item with the lowest communality until a solution with five items was found namely a, g, i, q and s. The data was found to be unidimensional and the scree plot is shown in Figure 2.

All items had communalities ranging from .405 to .662 which, while less than the .5 discussed earlier was considered adequate. The Kaiser-Meyer-Olkin measure was .834 and Bartlett's test of sphericity ($\chi 2$ (10) = 348.7, p<.001) reached statistical significance, supporting the factorability of the correlation matrix. The factor explained 48.4% of the variance. Table 4 shows the factor matrix for the EFA. The items that cluster on the factor suggest that this does represent sexual discrimination. All loadings were above .6. It was decided to retain the five items for testing in the confirmatory factor analysis (CFA).

Table 4: Summary of Exploratory Factor Analysis (EFA) results

	Factor 1
s. Considered for promotional opportunities.	.814
a. high organisational support	.711
i. successful in obtaining fair remuneration	.651
q. allocated roles with lower levels of responsibility (rev)	.651
g. “fit in” well	.637

4.3 Confirmatory Factor Analysis (CFA)

The purpose of the confirmatory factor analysis (CFA) in scale development is to establish the validity of a scale following an exploratory factor analysis (EFA) (De Vellis, 2003). Worthington and Whittaker (2006) note is now common practice for researchers to use Structural Equation Modelling (SEM) as the preferred approach for CFA. An independent data sample of 239 participants was used for the CFA. This sample size was considered ample for this analysis. Kline (2005) recommends a minimum sample of between 100 and 200 for SEM and Grimm and Yarnold (1995) recommend between five and 10 participants per observed variable. Researchers typically use a Chi-Square test statistic as a test of overall model fit in SEM however other fit indices, which may be classified as incremental, absolute or predictive, may also be used to evaluate model fit (Kline, 2005). Based on the EFA a one-factor model of sexual discrimination was tested. (*Please note that item q was denoted with an “r” to indicate it had been reverse coded as required*.)

A CFA using Maximum Likelihood (ML) estimation on the co-variance matrix of five items thought to represent sexual discrimination found that the model was an excellent fit to the hypothesized single-factor model $\chi 2$ (N=239, df = 5) = 5.776, p = .329. Another measures of close fit was the RMSEA = .02 (.000; .09). The five items reflecting Sexual Discrimination were Organisational support (a), “Fit in” well (g), Promotional opportunities (s), Fair remuneration (i) and Role responsibility (qr). The table of regression weights (refer Table 5) refers to the unstandardised parameter estimates for the factors loadings. No critical ratios (t-values) are stated for the factor loading of (a) because it was fixed to unity. All remaining factor loadings were significant.

The factor loadings were significant at p = .001, and standardised loadings ranged from .60 to .92 (refer Table 5). This is reflected in Figure 3 which also shows the standardised loadings.

The reliability of the sexual discrimination measure was checked using the whole sample and the co-efficient Alpha was found to be .81. The deletion of any of the items resulted would have resulted in a lower value thus again justifying the retention of the five items. This value of .81 is well above the minimum of .7 (De Vellis, 2003).

Table 5: Regression Weights for Model

			Estimate	S.E.	C.R.	*P*
a	<---	SEXUAL DISCRIMINATION	1.000			
g	<---	SEXUAL DISCRIMINATION	.917	.100	9.159	***
s	<---	SEXUAL DISCRIMINATION	.896	.084	10.669	***
i	<---	SEXUAL DISCRIMINATION	.777	.092	8.478	***
qr	<---	SEXUAL DISCRIMINATION	.598	.079	7.593	***

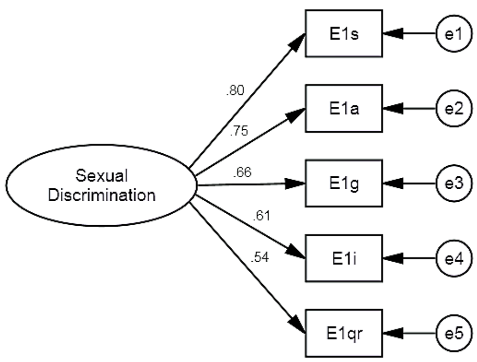

Figure 3: Standardised parameter estimates for sexual discrimination scale

5. Discussion and conclusions

The scale was subject to tests for normality, an EFA and finally a CFA conducted. Theoretically the scale was expected to be bidimensional, with a factor representing sexual discrimination and one representing sexual harassment. The trial scale was reduced from 20 to 11-items in order for the data to be considered to be reasonably "normal" and meet the criteria for continuous data. This also resulted in the loss of the two items related to sexual harassment. The remaining 11-items were then subjected to an EFA using common-factors analysis (principal-axis factoring) with oblique rotation (Direct Oblimin). Due to low communalities 6 items required removal. The CFA found that the model was an excellent fit to the hypothesised single-factor model. The final Sexual Discrimination measure had 5-items namely "Organisational support (a)", "Fit in" well (g)", "Promotional opportunities (s)", Fair remuneration (i) and Role responsibility (qr). The factor explained 48.4% of the variance and the Cronbach Alpha for the scale was excellent at .81. As scale development can include 3 to 4 times the number of final items in the measure (Worthington and Whittaker, 2006) this reduction was considered acceptable.

The study included a number of lessons which are summarised below. The author wishes she had known these before starting the study and therefore hopes it assists other researchers!

1 Scale development is more complex than it may seem initially. At least two independent samples (which need to be quite large) are needed to validate a new measure before it is used in subsequent studies.
2 Researchers should ensure that there as an absolute minimum 3-items/expected factor.
3 More than a 5-point response format is a good idea. This study could have used a 11-point response (5 either side of neutral point – see below) or at least a 7-point format.
4 While it was admirable that the organisational psychologist academics considered gender equity as the most likely response, this was not the case and the responses were higly skewed. The 11-point response (noted above) would have allowed for transformation of data - if required. Alternatively a 7-point format may have been more appropriate if an agreement type response had been used. Also a mixed gender sample may have reduced the skewness.
5 Reverse scoring of items is problematic, with most not making the final measure. This is an issue which more recently has been raised within scale development circles due to the greater sensitivity of SEM. Many researchers now recommend no reversed items!

References

Babakus, E., Ferguson, C.E., & Jöreskog, K.G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research*, *37*, 72-141.

Bergman, B., & Hallberg, L.R.M. (2002). Women in a male-dominated industry: Factor analysis of a women workplace culture questionnaire based on a grounded theory model. *Sex Roles*, *46*(9-10), 311-322.

Bollen, K.A. (1989). *Structural equations with latent variables*. New York, NY: Wiley.

Comrey, A.L. (1973). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Dainty, A.R.J., Bagilhole, B.M., & Neale, R.H. (2001). Male and female perspectives on equality measures for the UK construction sector. *Women in Management Review, 16*(5/6), 297-304.

De Vellis, R.F. (2003). *Scale development: Theory and application*. Applied Social Research Methods Series, Volume 26, 2nd edition, Thousand Oaks, CA: Sage Publications.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London, UK: Sage.

Finney, S.J., & DiStefano, C. (2006). Non-normal and categorical data. In G.R. Hancock, & R.O. Mueller (Eds.), *Structural equation modeling: A second course*. Charlotte, NC: Information Age.

Francis, V (2015) The development of a scale to measure gender equity in construction, *Eighth International Conference on Construction in the 21st Century (CITC-VIII).* May 27-30, 2015, Thessaloniki, Greece.

Gorsuch, R.L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Grimm, L.G., & Yarnold, P.R. (1995). *Reading and understanding multivariate statistics*. Washington, DC: American Psychological Association.Guadagnoli and Velicer (1988)

Guadagnoli, E., & Velicer, W.F. (1988). Relation to sample size to the stability of component patterns. *Psychological Bulletin*, *103*(2), 265-275.

Gale, A.W. (1994). Women in non-traditional occupations: The construction industry. *Women in Management Review*, *9*(2), 3-14.

Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist*. London, UK: Sage.

Kaiser, H.F. (1958). The varimax criterion for analytical rotation in factor analysis. *Psychometrika*, *23*(3), 187-200.

Kline, R.B. (2005). *Principles and practice of structural equation modelling* (2nd ed.). New York, NY: Guilford.

Loehlin, J.C. (1998). Latent variable models: An introduction to factor, path and structural analysis, 3rd edition. Mahwah, NJ: Lawrence Erlbaum.

Muthén, B., & Kaplan, D. (1985) A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology Society*, *38*, 171-189.

Tabachnick, B.G., & Fidell, L.S. (2007). *Using Multivariate Statistics* (5th ed.). Boston, MA: Allyn & Bacon.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington DC: American Psychological Association.

Velicer, W.F., & Fava, J.L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, *3*(2), 231-251.

Watts, J.H. (2007). Porn, pride and pessimism: Experiences of women working in professional construction roles. *Work, Employment and Society*, *21*(2), 299-316.

Worthington, R.L., & Whittaker, T.A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*(6), 806-838.